IN THE CLAIMS

Please amend the claims as follows:

1.   (Currently Amended) A server comprising:

a processor;

a memory;

a system area network connection;

a local area network connection;

wherein the processor, memory, system area network connection, and local area network connection are operably interconnected within the ~~system~~ server; and

software held in the memory and operable on the processor to:

load unique content into the memory from a storage location,

receive requests for content over the local area network connection,

service requests for the content in memory,

service requests for content located in a <u>random access</u> memory of another server <u>prime with content</u> by obtaining the content over the system area network connection, wherein the another server is identified as a function of a table holding content availability and location data of content <u>primed to and</u> held in <u>random access memories</u> ~~memory~~ of one or more other servers; and

cache content used to service request for content located in the memory of the other server for use in servicing subsequent requests for identical content.

2.   (Original) The server of claim 1, wherein the software operable on the processor is a component of an operating system of the server.

3.   (Original) The server of claim 1, wherein the software operable on the processor is a driver.

4.   (Original) The server of claim 1, wherein the software operable on the processor is a middleware component.

5.   (Original) The server of claim 1, wherein the system area network is a Gigabit Ethernet network.

6.   (Original) The server of claim 5, wherein the Gigabit Ethernet network is TCP Offload Engine enabled.

7.   (Original) The server of claim 1, wherein the unique content is loaded into memory prior to the server being available to service content requests.

8-14.   (Canceled)

15.   (Currently Amended) A method of server operation comprising:

        priming a random access memory of a server, wherein the server is a member of a server cluster, wherein the content in the random access memory of the server is unique to the server amongst all servers in the server cluster;

        making the content in the server random access memory available to other servers in the server cluster over a high-speed interconnection;

        receiving requests for content;

        fulfilling content requests by retrieving data from the server random access memory and from random access memories of one or more other servers over the high-speed interconnection; and

        caching content of other servers that has been requested either recently or commonly to provide the server the ability to fulfill requests for cached content locally.

16.   (Original) The method of claim 15, wherein requests for content are received over a local area network connection.

17.   (Original) The method of claim 16, wherein requests are received into the local area network on a router coupled to the Internet.

18-23. (Canceled)